

# GHAJAR EXHIBIT 25

UNITED STATES DISTRICT COURT  
NORTHERN DISTRICT OF CALIFORNIA  
SAN FRANCISCO DIVISION

The discovery deposition of DANIEL SPULBER, taken in the above-entitled cause, before Deborah A. Thompson, a Certified Shorthand Reporter in the State of Illinois, taken pursuant to the Federal Rules of Civil Procedure of the United States District Courts at 10 North Dearborn Street, Chicago, Illinois, on the 3rd day of March, 2025, at 9:30 o'clock a.m.

1 that.

2 Q Let's talk about that segment, the segment  
3 of AI training. When you talk about licensing,  
4 licensing training data, right, is that segment both  
5 for copyrighted and non-copyrighted material or only  
6 copyrighted?

7 A I'm not sure. I think there may be some  
8 data that's licensed that may not be protected by  
9 copyright. I really don't know. I think my focus is  
10 on licensing of copyrighted books. But it's possible  
11 that I'm paying a license to use data that is subject  
12 to other protections and ownership and so on. I  
13 can't tell you every possibility. So there may be  
14 some that's not.

15 Q So in your market segment of licensing for  
16 AI training, the agreements in that segment could be  
17 for material that has no IP protection at all?

18 A I didn't say that. I said it may be  
19 subject to other kinds of IP protection.

20 Q So let me ask the question directly. For  
21 agreements governing data that has no IP protection  
22 at all, do you consider that to be part of the market  
23 for licensing data for training an LLM?

24 MS. GEMAN: Objection.

1                   THE WITNESS: So this is a vast market, and  
2 it includes all sorts of stuff, including stuff  
3 that's not copyrighted. So there may be exchange of  
4 data under some terms or others that don't look like  
5 traditional licenses. So I would prefer to be  
6 inclusive rather than not because there's so much --  
7 so many transactions in this marketplace. But by and  
8 large, I would expect that licenses would be for  
9 copyrighted material, but there could be other kinds  
10 of data that's protected, for example, by trade  
11 secrets and so on.

12 BY MR. BRIGHAM:

13 Q       Right, they could be paying for a license  
14 other than for copyrighted material because the  
15 market is so broad?

16 A       Yes. They could be paying to license  
17 material that's not necessarily protected by  
18 copyright. There might be other protections like  
19 privacy. For example, trade-secret type data or data  
20 that's been produced as a work product, this kind of  
21 thing.

22 Q       Or there may be no protections at all;  
23 right? They might be just paying for access out of  
24 convenience?

1 A That's also possible.

2 Q So if we go back to your opening report,  
3 page 14.

4 A Yes.

5 Q That Table 1 -- there's a number of open AI  
6 agreements on here. Do you see those?

7 A I see some of them, yes. There's a lot of  
8 them here.

9 Q Do you recall reviewing any of those  
10 agreements?

11 A I don't recall. I may have. I don't  
12 recall.

13 Q You mentioned collective agreements before  
14 or collective arrangements. Are you relying on the  
15 details of any existing collective arrangements as  
16 somewhere that Meta could go license plaintiffs'  
17 works right now? Or are you just offering the  
18 opinion that they could exist in the future?

19 MS. GEMAN: Objection.

20 THE WITNESS: So there are various  
21 collective rights organizations. I list many of  
22 them. And one particular one is very important, and  
23 that's the CCC, the Copyright Clearance Center. They  
24 have experience in collective rights management.

1 copies also wasn't relevant to your opinion; right?

2 A I believe that making a copy is harm.

3 Q But you don't know whether they lost any  
4 sales --

5 MS. GEMAN: Objection.

6 BY MR. BRIGHAM:

7 Q -- as a result of the copy; right?

8 A It's not possible to quantify that effect.  
9 They lost the sales of the copy. In other words, if  
10 Meta made one copy or two or a million, those are  
11 lost sales.

12 Q And those are the only lost sales that you  
13 believe occurred?

14 A No. If, in addition, Meta's products lead  
15 to diminished usage and sales of their work by  
16 providing substitutes, that's another loss of sales.

17 Q So we'll get to that. But you have not  
18 looked at plaintiffs' sales numbers?

19 A That is correct.

20 Q You haven't looked at them; right?

21 A That's correct.

22 Q And so therefore you have not determined  
23 whether there was a drop in sales or lost sales after  
24 the alleged copying; correct?

1 MS. GEMAN: Objection. Misstates  
2 testimony.

3 THE WITNESS: So I have not looked at  
4 whether plaintiffs had lost sales other than the fact  
5 that, when Meta made copies, they lost sales. I can  
6 tell you that making copies of copyrighted works and  
7 putting them online through torrent sites or  
8 downloading or making copies, that is the essence of  
9 lost sales.

10 BY MR. BRIGHAM:

11 Q In your opinion, Llama wouldn't lead to the  
12 discovery of plaintiffs' works; right?

13 MS. GEMAN: Objection.

14 THE WITNESS: I don't know.

15 BY MR. BRIGHAM:

16 Q You don't have an opinion one way or the  
17 other whether Llama would lead to the discovery of  
18 plaintiffs' works?

19 A Could you explain what you mean by "the  
20 discovery of plaintiffs' works"? I'm not sure I  
21 understand your question.

22 Q If I use Llama asking questions about  
23 books, do you believe that Llama could be beneficial  
24 and that it will direct traffic toward plaintiffs'

1                   So my impression is, as I've been asked to  
 2 assume and take as a given, that Meta is infringing  
 3 on plaintiffs' works, and that infringement is not  
 4 designed in any way to benefit the authors, but  
 5 rather it's designed to create benefits to Meta.

6 BY MR. BRIGHAM:

7                   Q     Have you personally seen any instances  
 8 where the Meta LLM outputted correct information  
 9 about plaintiffs' books?

10                  A     I have not looked at output of Meta.

11                  Q     So you don't know whether output of Meta  
 12 regurgitates parts of books?

13                  MS. GEMAN: Objection.

14                  THE WITNESS: I don't know. It's possible,  
 15 but this is for experts.

16 BY MR. BRIGHAM:

17                  Q     That's not part of your opinion?

18                  A     No. It is possible. And it's something  
 19 that would be factored into the creation of economic  
 20 substitutes for authors' works. So I can't provide  
 21 evidence for that because that's not my job. My job  
 22 is to provide an economic analysis of this. But if  
 23 indeed it outputs something about authors' works, it  
 24 may create substitutes. Or it can harm authors'

1                   reputations as I explain in my statement.

2                   Q         But that's hypothetical; right? You have  
3                   no proof that infringing content was actually  
4                   generated by a Llama model?

5                   MS. GEMAN: Objection.

6                   THE WITNESS: This is beyond my analysis,  
7                   but there is certainly a risk -- and that's not  
8                   beyond my analysis. There's a risk that it may  
9                   output content that affect's authors' sales.

10                  BY MR. BRIGHAM:

11                  Q         I understand why you think it may, but you  
12                  have no proof that it has?

13                  MS. GEMAN: Objection.

14                  THE WITNESS: That's correct.

15                  BY MR. BRIGHAM:

16                  Q         So for the third type of harm,  
17                  substitution, how many sales were lost for each  
18                  plaintiff due to substitution by LLM generated works?  
19                  You didn't calculate that; right?

20                  MS. GEMAN: Objection. Misstates  
21                  testimony.

22                  THE WITNESS: So I did not calculate lost  
23                  sales, but I provide evidence that leads me to  
24                  believe that there my be lost sales.

1 BY MR. BRIGHAM:

2 Q But you don't know how many sales, if any,  
3 were lost due to substitution?

4 A I believe I've answered that question.

5 Q And the answer was you don't know?

6 MS. GEMAN: Objection. Misstates  
7 testimony --

8 BY MR. BRIGHAM:

9 Q Is the answer "I don't know"?

10 MS. GEMAN: Sorry. I wasn't done with my  
11 objection. Misstates testimony about the meaning of  
12 lost sales.

13 THE WITNESS: So I've mentioned lost sales  
14 in terms of the fact that Meta made copies. I've  
15 mentioned the risk of substitution or economic  
16 substitution due to the output of Meta. But I have  
17 not quantified those amounts.

18 BY MR. BRIGHAM:

19 Q And you're not aware of any instance of  
20 substitution?

21 A You have to flesh that out a little bit.

22 Q You're not aware of any instance where an  
23 output generated by Meta AI acted as a substitute for  
24 one of plaintiffs' books?

1 A No, I am not.

2 Q Did you compare the sales of each book of  
3 plaintiffs before introduction of a Llama model and  
4 then after?

5 A No, I did not.

6 Q Why not?

7 A It wasn't necessary for my analysis.

8 Q You could have, though?

9 MS. GEMAN: Objection.

10 THE WITNESS: I'm not sure that information  
11 is available.

12 BY MR. BRIGHAM:

13 Q Adequate information from plaintiffs was  
14 not available?

15 MS. GEMAN: Objection.

16 THE WITNESS: To fully do this, I need a  
17 little bit longer time horizon, and I would need  
18 sales data for other books and so on to make a  
19 complete analysis. You can't just look at one or two  
20 or a few books. What you can do is you can say that,  
21 you know, licensing didn't occur. Authors were  
22 deprived of licenses from sales from absence of  
23 licensing in other words. Or I can say that many  
24 copies were made, and authors were deprived of

1 order.

2 MR. LAUTER: In the protective order.

3 MR. STEIN: Because open AI, in making its  
4 production, expressed -- and I'll say this on the  
5 record -- expressed concerned that anyone but outside  
6 counsel would see their documents. And we told them  
7 that the protective order is what it is. So I just  
8 wanted to see if there had been any further  
9 discussions between you and AI to this effect.

10 MR. LAUTER: We're just following the  
11 protective order.

12 MR. STEIN: Okay. Thank you.

13 THE VIDEOGRAPHER: Off the record at 5:49.

14 (Recess taken.)

15 THE VIDEOGRAPHER: Back on the record at  
16 6:05.

17 BY MR. BRIGHAM:

18 Q Have you ever seen a book generated by an  
19 LLM?

20 MS. GEMAN: Asked and answered.

21 THE WITNESS: I believe I indicated to you  
22 that I don't know. I don't think so, but I don't  
23 know.

24

1 BY MR. BRIGHAM:

2 Q Do you know how many books generated by an  
3 LLM exist?

4 A No, I don't know.

5 Q Do you know how many books generated by an  
6 LLM have been sold?

7 A I don't know.

8 Q Are you aware of a single instance in which  
9 someone used Llama instead of reading one of  
10 plaintiffs' books?

11 A I don't know.

12 Q Are you aware of a single instance in which  
13 someone read a Llama output instead of reading one of  
14 plaintiffs' books?

15 A I don't know, but it doesn't affect my  
16 opinion.

17 Q Are you aware of a single instance in which  
18 someone wrote a book using Llama?

19 MS. GEMAN: Objection.

20 THE WITNESS: I am aware of people using AI  
21 LLMs to write books, and I've read reports that  
22 there's a flood of them. But I don't know which, if  
23 any, have been done with Llama.

24

1 BY MR. BRIGHAM:

2 Q So you can't tell me of a single instance  
3 in which someone wrote a single book using Llama?

4 A Again, I've read reports that there's a  
5 flood of books written with AI LLMs, but I don't know  
6 if any of those are from Llama.

7 Q Are you aware of a single instance in which  
8 someone used any AI model instead of reading the  
9 plaintiffs' books?

10 MS. GEMAN: Objection.

11 THE WITNESS: Again, I don't know. And  
12 you're just referring to plaintiffs' books. I don't  
13 know.

14 BY MR. BRIGHAM:

15 Q Are you aware of a single instance in which  
16 someone read any AI model output instead of reading  
17 plaintiffs' books?

18 A I don't know. I'm not aware.

19 BY MR. BRIGHAM:

20 Q When you say that authors face the risk  
21 that Llama outputs could substitute for plaintiffs'  
22 books, you're speculating or guessing about something  
23 that might occur; right?

24 MS. GEMAN: Objection.

1                   THE WITNESS: I'm examining the risk of  
2 substitutes based on my observation of AI LLM models  
3 generally.

4 BY MR. BRIGHAM:

5 Q       But you're saying it might occur in the  
6 future, fair?

7 A       Yes. I think that's one interpretation of  
8 the word "risk."

9 Q       And it might not occur; right?

10 A      That's possible as well.

11 Q       Okay. You talk about Meta's market  
12 capitalization in your report. Do you recall that?

13 A      Yes, I do.

14 Q       Did Meta's market cap impact the damages  
15 amount that you calculated?

16 A      No, it did not. And as you know, I  
17 calculated the damages based on the market price of  
18 books used as LLM training data.

19 Q       You're not opining that Llama was the cause  
20 of the increase in the market cap, are you?

21 A      No. I'm not making a causation statement,  
22 but I am pointing out that during the period that  
23 Meta developed its AI LLM models, its market  
24 capitalization rose by over a trillion dollars, I

1 right?

2 A So these types of projections have to do  
3 with revenue.

4 Q So you don't believe that anyone would  
5 benefit from the trillions of dollars?

6 MS. GEMAN: Objection.

7 THE WITNESS: I never said that.

8 BY MR. BRIGHAM:

9 Q Can the growth -- you point to this growth  
10 to trillions of dollars. Can that growth be achieved  
11 without the use of plaintiffs' books in training?

12 MS. GEMAN: Objection.

13 THE WITNESS: I believe that the training  
14 and development of AI models, such as Meta's Llama  
15 models, requires extraordinarily large amounts of  
16 data, and there are reports that more and more data  
17 will be required in the future. For example,  
18 companies have suggested that they need nuclear power  
19 plants to be established to run their computers.  
20 That will be used to process and compute with that  
21 data that's being fed into these artificial  
22 intelligence models. So I believe that a lot more  
23 data is going to be required.

24 Now, of that data, I believe that the

1 STATE OF ILLINOIS )  
2 ) SS:  
3 COUNTY OF C O O K )

4 I, Deborah A. Thompson, a Certified  
5 Shorthand Reporter within and for the State of  
6 Illinois, do hereby certify:

7 That previous to the commencement of the  
8 examination of the witness, the witness was duly  
9 sworn to testify the whole truth concerning the  
10 matters herein;

11 That the foregoing deposition was reported  
12 stenographically by me, was thereafter reduced to a  
13 printed transcript by me, and constitutes a true  
14 record of the testimony given and the proceedings  
15 had;

16 That the said deposition was taken before  
17 me at the time and place specified;

18 That the reading and signing by the  
19 witness of the deposition transcript was agreed upon  
20 as stated herein;

21 That I am not a relative or employee or  
22 attorney or counsel, nor a relative or employee of  
23 such attorney or counsel for any of the parties  
24 hereto, nor interested directly or indirectly in the  
outcome of this action.

25 IN WITNESS WHEREOF, I do hereunto set my  
26 hand at Chicago, Illinois, this 10th day of March,  
27 2025.

28 

29 Certified Shorthand Reporter  
30 State of Illinois  
31 CSR License No. 084-003487